

De la tesis doctoral de Tiselius a la proteómica: setenta y cinco años de electroforesis de proteínas

Juan Fernández Santarén

Arbor CLXXVII, 698 (Febrero 2004), 259-284 pp.

El término «proteoma» se utilizó por primera vez en 1995 para describir las proteínas codificadas por un genoma. De forma involuntaria el proteoma se ha ido transformando en «proteómica», un conjunto de técnicas al que se ha elevado al rango de disciplina científica. ¿Qué es la proteómica? En esencia es el estudio, a gran escala, de la composición y propiedades de las proteínas que integran un sistema biológico. Su ejecución práctica se basa fundamentalmente en dos pilares. En primer lugar la separación de las mezclas complejas de proteínas, proceso que habitualmente se lleva a cabo mediante electroforesis bidimensional y en segundo lugar la identificación de las proteínas separadas para lo que actualmente se recurre a la espectrometría de masas. El objetivo de este artículo es dar una breve noción de esas técnicas después de hacer una simbólica revisión histórica de los métodos de electroforesis de proteínas, desde las primitivas separaciones hasta las refinadas técnicas actuales.

Introducción

Desde que se publicó la secuencia del primer genoma completo de *Haemophilus influenzae*, se puso de manifiesto que muchas de las posibles proteínas codificadas por sus genes no tenían una función conocida y de aquellas con supuesta función, una gran parte habían sido atribuidas solo por analogía. Esta observación se ha ido repitiendo a medida que los

sucesivos genomas han ido viendo la luz y quizá ha alcanzado su punto culminante al completarse el proyecto de secuenciación del genoma humano, logro que ha marcado un hito en el desarrollo de la biología y que ha reformulado de forma sustancial algunos de sus principios. Se ha demostrado la posibilidad de acceder a la totalidad de las secuencias de los genes de un organismo, por muy complejo que sea, pero simultáneamente se ha puesto de manifiesto la necesidad de dar sentido a la enorme cantidad de datos que el propio genoma ha generando.

Ahora que se pueden visualizar genomas completos, la primera pregunta obvia es ¿cuáles son y cómo interaccionan sus productos proteicos?. Hay múltiples razones para centrarse en el análisis de las proteínas: por ejemplo el nivel de expresión del mRNA frecuentemente no se correlaciona con la cantidad de proteína activa en una célula y la secuencia del gen no describe las posibles modificaciones post-traduccionales que pueden ser esenciales para la función y actividad de una proteína. Es decir, el estudio del genoma no da cuenta de la dinámica de los procesos celulares.

Por tanto, la investigación ha fijado como meta la posibilidad de identificar y analizar la estructura e interacciones de las diferentes proteínas codificadas por los genes individuales con la esperanza de conocer en detalle su función. Ahora se empieza a cambiar la interpretación y significado del código de DNA volviendo la vista nuevamente hacia las proteínas, complejos proteicos y ensamblajes celulares. Este intento de estudio global de las proteínas de un sistema biológico es lo que se ha denominado proteómica y se ha considerado como vía fundamental y lógica de la investigación en la llamada era «post-genómica». La proteómica ocupa una posición central en la nueva biología. Esta disciplina, que ha surgido a partir del laborioso trabajo de décadas de análisis de proteínas en geles bidimensionales se ha revitalizado por el desarrollo de la espectrometría de masas y el crecimiento de los bancos de datos de secuencias en los que realizar búsquedas, posibilitando los estudios de expresión diferencial a nivel de proteínas.

Por tanto, y resumiendo, la proteómica, en su conjunto utiliza una combinación de sofisticadas técnicas que en una versión simplificada podríamos decir que incluyen la electroforesis bidimensional, el análisis de las imágenes de los geles y la espectrometría de masas que unida a la bioinformática trata de resolver, cuantificar y caracterizar las proteínas de un sistema biológico. De la aplicación de la proteómica cabe esperar igualmente una visión integrada de los procesos patológicos individuales a nivel de proteínas. En cualquier caso, sus orígenes hay que buscarlos en la separación electroforética de proteínas, una técnica con mas de setenta y cinco años de andadura.

Algunos datos históricos sobre la electroforesis de proteínas

La electroforesis podríamos definirla como la técnica de separación de partículas en base a sus diferente comportamiento ante la aplicación de un campo eléctrico. La técnica se ha ido diversificando y sofisticando de manera progresiva hasta convertirse en imprescindible dentro de la actual biología molecular en la separación analítica o preparativa de proteínas y ácidos nucleicos.

Fijar los comienzos históricos de un proceso científico no es, en ocasiones, una tarea fácil ya que no esta exenta de cierta arbitrariedad cuando no de error. Tal es el caso de la electroforesis. ¿donde marcamos su comienzo?. Posiblemente si quisiéramos ser rigurosos habría que remontarse hasta 1791, momento en el que Faraday enuncia las leyes de la electrolisis, o a 1809 cuando encontramos los primeros experimentos de electroforesis llevados a cabo por el investigador ruso Reuss. En cualquier caso, resulta obligado mencionar que el siglo XIX fue testigo de una serie de lentas aportaciones teóricas y prácticas que sentaron las bases para el posterior desarrollo de la técnica. Citemos, por ejemplo, que fue entonces cuando se hicieron las primeras observaciones resultantes de aplicar un campo eléctrico en los extremos de un capilar de vidrio horizontal relleno de una solución salina que contenía partículas cargadas. Se comprobó que la superficie interna de la pared del tubo adquiría carga negativa mientras que la capa de solvente que rodeaba dicha pared adoptaba carga de signo opuesto, hecho que ya había sido postulado por Helmholtz. El movimiento resultante del líquido próximo a la pared hacia el ánodo se denominó electroendósmosis y pudo describirse por una ecuación.

Posteriormente los experimentos de Hittorf, Nernst y Kohlrausch permitieron determinar la conductancia y el radio de diversos iones de pequeño tamaño en solución, junto con sus números de transporte, es decir la fracción de corriente que es transportada por el ión.

Igualmente fundamental para el posterior desarrollo de la isotacoforesis fue la formulación de la función de Kohlrausch que describía el orden de migración electroforética de una serie de iones y sus concentraciones relativas. Este autor fue capaz de describir el proceso y formular las ecuaciones que gobiernan el comportamiento de un campo eléctrico en una interfase, es decir en el estrecho frente que definen una serie de partículas coloidales en movimiento. La aportación resulto de gran utilidad años mas tarde para el análisis de iones de mayor tamaño y especialmente de proteínas. A la vuelta del siglo XX se vio que era posible obtener frentes de movimiento bien definidos para especies ionizadas si se

utilizaba un tubo de vidrio en forma de U en cuyo interior se aplicaba un campo eléctrico. Obligado resulta recordar en este punto que Arrhenius vio reconocida su aportación al conocimiento de los iones en soluciones

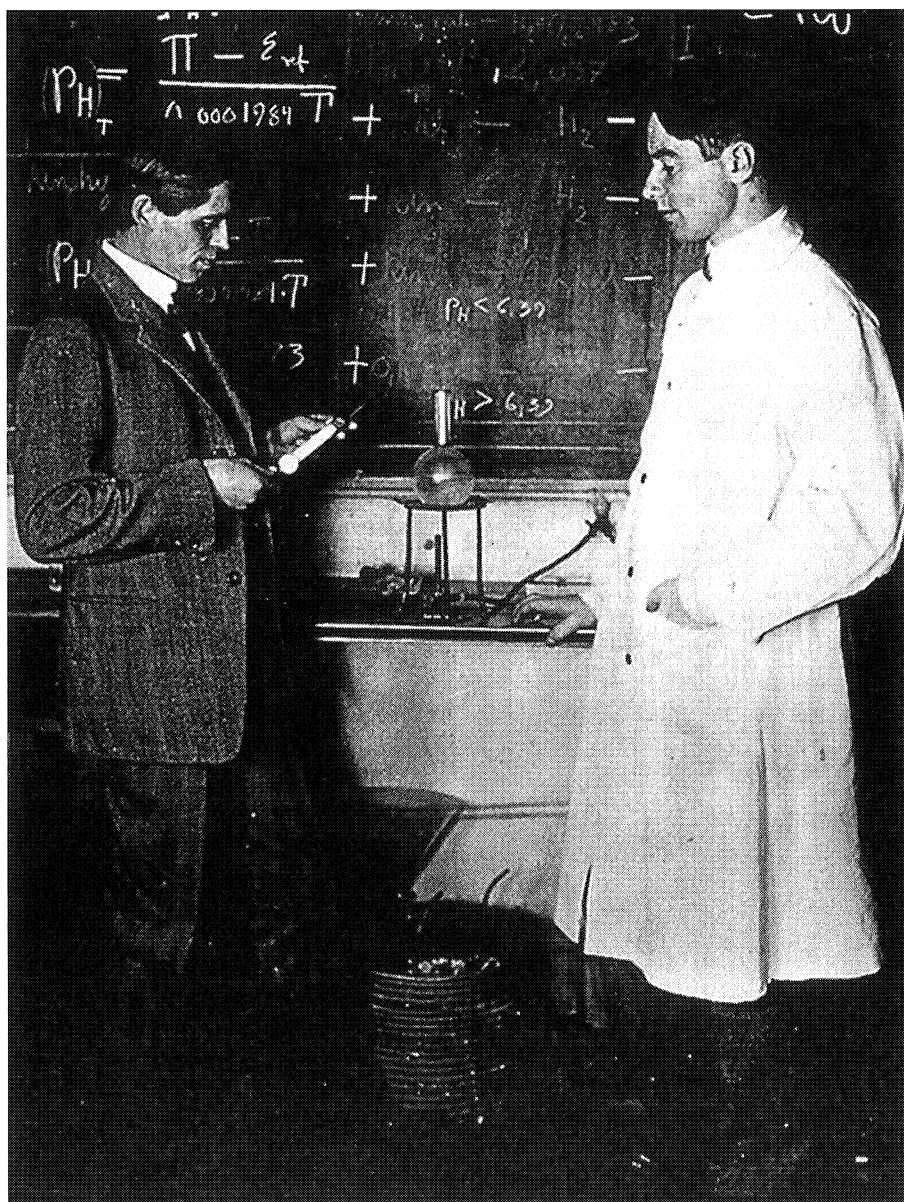


FIGURA 1. Theodor Svedberg y su estudiante Arne Tiselius en el laboratorio de Estocolmo alrededor de 1926.

acuosas mediante su teoría de la disociación al recibir el Premio Nóbel en 1903.

Hecho este breve preámbulo histórico, aproximémonos ya al punto que da título al artículo. Debemos situarnos para ello en el Instituto Karolinska de Estocolmo, en el laboratorio de Theodor Svedberg. El año podemos fijarlo en 1926, el mismo en el que el investigador sueco era galardonado con el Premio Nóbel por sus trabajos sobre proteínas, fundamentalmente mediante técnicas de ultracentrifugación. En esa época Svedberg propuso a un joven estudiante de su laboratorio, Arne Wilhelm Kaurin Tiselius trabajar en electroforesis de proteínas analizando su migración mediante la técnica de frente móvil, utilizando sistemas ópticos de detección análogos a los que a él le habían proporcionado tantos éxitos en centrifugación (Figura 1). En 1930 Tiselius publicó su Tesis Doctoral: «El método del frente móvil en el estudio de electroforesis de proteínas» en la que presentaba una nueva técnica para examinar algunas propiedades químico físicas de estas macromoléculas. Los experimentos los llevó a cabo en un tubo de cuarzo en forma de U en cuyo interior las proteínas ionizadas movían libremente y sus frentes de avance se detectaban y fotografiaban utilizando luz ultravioleta, siguiendo las normas de las antiguas observaciones de Toepler, cuando publicó el método de Schlieren. Normalmente, con el uso de este principio, solo el componente que migrase más rápido podría ser aislado. Uno de los problemas que tenía la técnica era la habitual pérdida de resolución de los frentes de proteína debido a la convección térmica causada por el calentamiento eléctrico de la solución. En 1937 Tiselius redujo este efecto introduciendo una celda electroforética con una sección rectangular en un recipiente a 4 °C, donde el agua tiene su máxima densidad. La mejora hizo de la electroforesis de frente móvil un método analítico de cierta precisión. Así, Tiselius fue capaz de describir cuatro frentes móviles en el análisis de suero correspondientes a la albúmina y a la α - β - y γ -globulina. Además, vio que esta última aumentaba después de infección e inmunización. En 1948 Tiselius obtuvo el Premio Nóbel por su contribución al desarrollo del método del frente móvil y sus aportaciones en el análisis de la cromatografía de adsorción.

En 1938 Philop publicó el método de la cámara astigmática de Schlieren que permitía el registro directo sobre una película fotográfica de las variaciones de los índices de refracción, que eran proporcionales a los gradientes de concentración de las proteínas en el interior de la celda. Este equipo de electroforesis de frente móvil, desarrollado en sus inicios por Tiselius, fue comercializado más tarde por LKB y otras compañías (Figura 2).

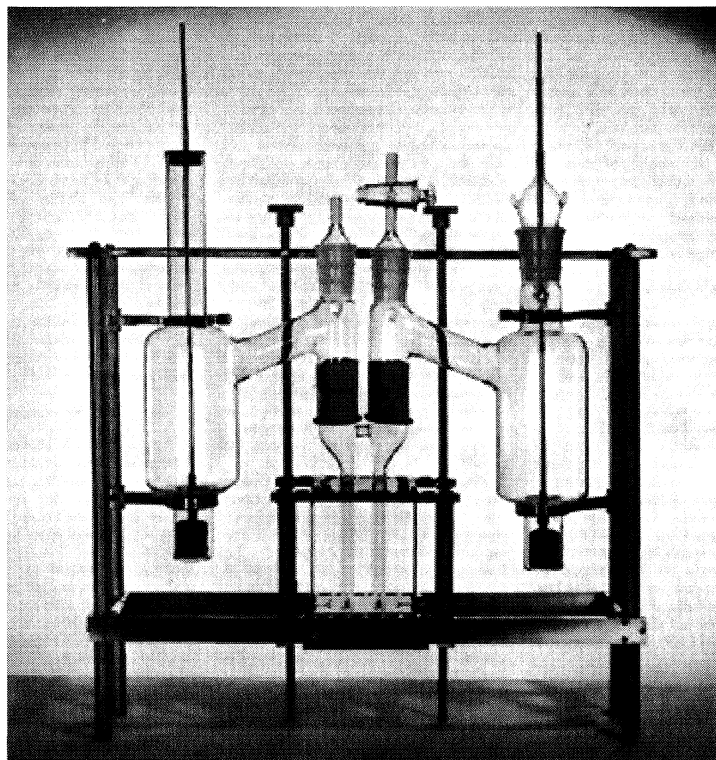


FIGURA 2. Equipo de electroforesis libre utilizado en la época de Tiselius. El sistema tiene un tubo en forma de U en la parte inferior central. Los electrodos están ubicados en los recipientes de vidrio de la derecha e izquierda. El tubo en forma de U era lo mas estrecho posible para facilitar su enfriamiento y prevenir las corrientes de convección.

Estos fueron los comienzos, meritorios y esenciales pero insuficientes. La electroforesis de frente móvil llegaba hasta un límite y si bien tenía un aceptable poder analítico para la época, adolecía de muy serias limitaciones. Durante la electroforesis existía una tendencia a que se mezclaran los componentes debido a que normalmente poseían una densidad superior a la del electrolito y porque la corriente eléctrica aplicada provocaba, al calentar la solución, alteraciones de convección térmica. Además, una vez desconectada la corriente, los componentes separados volvían a mezclarse dado que el medio en el que se encontraban era líquido y por tanto permitía su difusión. El problema trato de paliarse mediante ingeniosos diseños de los tubos en U a los que se doto de sofisticados mecanismos que permitían aislar una serie de compartimentos, pero a todas luces insuficientes.

En 1950 Haglund y Tiselius describieron un aparato preparativo en el que la electroforesis se llevaba a cabo en un tubo vertical lleno de esferitas de vidrio para contrarrestar la convección. El método fue mejorado por Flodin y Kupke sustituyendo las bolitas de vidrio por celulosa tratada con etanol-ClH. De esta forma se conseguía la estabilización frente a la convección suprimiendo muy eficazmente la electroendósmosis y la adsorción de las proteínas.

Era evidente que si se quería estabilizar los compuestos una vez separados mediante electroforesis esta se debía llevar a cabo sobre un soporte y no en condiciones de electroforesis libre. Había que desarrollar la denominada «electroforesis de zona» en la que los componentes migraran como «zonas» bien definidas sobre un soporte. Svensson, discípulo de Tiselius, hizo, junto con Longsworth y Dole, importantes contribuciones teóricas que resultaron fundamentales para el desarrollo de esta nueva modalidad de la técnica.

El precedente hay que buscarlo en 1923, año en que se utilizó un principio de electroforesis distinto del frente móvil. Kendall y Crittenden tuvieron éxito en la separación de metales de tierras raras y algunos ácidos simples por, según su denominación «el método de migración de ión», que era realmente isotacoforesis (del griego: iso = igual; taco = velocidad). Se postuló que los iones no solo se separaban, sino que también adaptaban sus concentraciones de acuerdo con la función de Kohlrausch. Kendall sugirió el uso de un ión coloreado con una movilidad intermedia entre los iones de interés. Posteriormente, dichos iones intermedios, que mejoraban la separación, se denominaron espaciadores. Kendall propuso igualmente algunos ingeniosos métodos de detección, como termometría, conductimetría y espectrometría, especialmente para el análisis de metales.

Por tanto era necesario resolver el problema de los soportes electroforéticos y desarrollar principios diferentes. Uno de las primeras alternativas a las que se recurrió, obvia por su sencillez, fue llevar a cabo la electroforesis en soluciones acuosas utilizando papel de filtro como medio soporte. Tuvo un gran éxito a partir de 1950 tras las contribuciones de Wieland y Fischer especialmente en estudios rutinarios de análisis clínicos de suero. Su buena acogida era lógica. En primer lugar, hasta ese momento, la electroforesis de frente móvil desarrollada por Tiselius era prácticamente el único equipo comercial disponible para llevar a cabo la técnica. Era caro y requería de gran espacio de laboratorio para ubicarlo, mientras que los requisitos para llevar a cabo una electroforesis sobre papel eran mínimos, el sistema era barato y bastaba un pequeño espacio para su instalación. Además reducía considerablemente la cantidad de muestra necesaria para llevar a cabo un análisis permitiendo su visuali-

zación una vez finalizada la separación mediante métodos selectivos de tinción. No tardó en diseñarse, por parte de Svenson, un equipo para llevar a cabo la electroforesis sobre papel que también comercializó LKB y alcanzo gran éxito y popularidad.

A pesar de sus evidentes ventajas sobre los métodos de electroforesis libre la electroforesis sobre papel también presentaba serias limitaciones. Escasa resolución, electroendósmosis, evaporación de electrolito y dificultades de tinción eran algunas de ellas. Eso hizo que se siguieran buscando soportes mas adecuados y que las miradas y los experimentos, se dirigieran hacia los geles. Y en este punto debemos mencionar que los geles de gelatina y agar ya se habían utilizado como soportes electroforéticos años atrás, pero su éxito y desarrollo había sido limitado por lo que en 1955 eran muy escasos los trabajos publicados al respecto. Uno de los primeros geles que se utilizó para la separación de proteínas fueron los constituidos por almidón, resultando esencial el trabajo de Smithies en este campo. Esos geles aumentaban significativamente la resolución alcanzada con el papel en la separación de las proteínas de suero. Además el gel podía ejercer un efecto de tamizado molecular que mejoraba considerablemente la resolución obtenida en condiciones de electroforesis libre. Pero seguían existiendo inconvenientes y uno no menor era la baja reproducibilidad de los experimentos debido a la imposibilidad de controlar el tamaño de poro en los geles, un parámetro fundamental en estas técnicas de separación.

Siguieron las pruebas y las aportaciones y es digno de mención el empleo como soporte, por parte de Flodin, de geles a base de dextrano granulado en electroforesis de zona. Este medio también lo utilizó Hjertén para separar electroforéticamente proteínas coloreadas y compuestos de pequeño tamaño molecular. Hjertén observó que los compuestos de bajo peso molecular se desplazaban mas lentamente, descubriendo de esta forma las propiedades de filtración molecular de los geles de dextrano (denominados posteriormente Sephadex) que tan fundamental resultó para el posterior desarrollo de las cromatografías. Los geles granulados de dextrano no eran soportes adecuados porque a la separación obtenida por la electroforesis se superponía el efecto de filtración antes mencionado que imposibilitaba la reproducibilidad de los resultados como en el caso del almidón. Resultaba esencial encontrar un soporte en el que poder controlar el tamaño de poro. El propio Hjertén observó que los geles a base de poli(acrilamida) podían cumplir dichos requisitos pero la fortuna no fue su aliada. Mientras que ultimaba sus experimentos vio como Raymond, Weintraub, Davis y Ornstein publicaban el uso de este polímero como soporte para electroforesis por lo que Hjertén no se molestó en publicar sus

resultados. No terminaron aquí sus aportaciones. Hjertén también tuvo éxito al adaptar la electroforesis analítica en gel con fines preparativos y en el desarrollo de métodos de localización de las proteínas en los geles de poliacrilamida basados en un barrido con luz ultravioleta a dos longitudes de onda diferentes, o cortando el segmento de gel que contenía la proteína de interés y eluyéndola mediante electroforesis en un gradiente de sacarosa-sal. No fue la acrilamida el único medio sobre el que trabajó Hjertén quien también desarrolló un método de electroforesis utilizando geles de muy baja concentración de agarosa (< 0.2 %) para el fraccionamiento de proteínas, ribosomas y virus.

En 1964 Ornstein y Davis introdujeron la electroforesis discontinua o «disc electroforesis». El uso de tampones con diferente pH e iones de distinta movilidad permitía concentrar las proteínas en bandas muy estrechas antes de iniciar el proceso de separación propiamente dicho. Se utilizaron geles de poliacrilamida como medio estabilizante que ejercía un importante efecto de tamizado molecular ya que las movilidades electroforéticas de las proteínas resultaban muy influenciadas al variar el tamaño de poro del gel. Para visualizar las proteínas en el gel después de la electroforesis se utilizó inicialmente Negro Amido, que fue rápidamente sustituido por el Azul Brillante de Coomassie, al ofrecer menores límites de detección para proteínas.

En 1967 Shapiro, Viñuela y Maizel publicaron un importante trabajo que describía la relación entre la movilidad electroforética de las proteínas en un gel homogéneo de poliacrilamida en presencia del detergente iónico dodecil sulfato sódico (SDS) y el radio molecular efectivo o, mas aproximadamente, el peso molecular. Comprobaron que la movilidad electroforética del complejo proteína-SDS era lineal, disminuyendo en función del logaritmo del peso molecular de las proteínas. Esta técnica se sigue utilizando actualmente para ensayar la homogeneidad y pureza de una proteína y fundamentalmente, para la determinación de su tamaño molecular en condiciones desnaturalizantes. En electroforesis en condiciones nativas, si se hacen estudios a diferentes concentraciones de gel, también es posible obtener una relación entre el coeficiente de retardo de una proteína y su peso molecular.

Un último refinamiento vino con el uso de geles contruidos con gradientes de acrilamida, es decir con un tamaño de poro decreciente hacia el ánodo. Aquí las proteínas migran hasta que quedan atrapadas en los poros del gel alcanzándose alta resolución. Con este tipo de geles en electroforesis discontinua y en condiciones desnaturalizantes es posible resolver del orden de 100 polipéptidos, la máxima capacidad de separación que presenta la electroforesis monodimensional.

El electroenfoque es la separación de compuestos sometidos a la acción de un campo eléctrico en un gradiente de pH. Los primeros intentos de esta técnica también se deben a Tiselius quien ya en 1941 trató de separar albúmina de suero y hemoglobina mediante «condensación isoelectrónica» utilizando un sofisticado aparato provisto de múltiples compartimentos y empleando sulfato sódico como electrolito. El dispositivo no funcionó de forma satisfactoria ya que como resultado de la electrolisis todos los iones de sal quedaban excluidos de la región próxima a pH 7 provocando una baja conductividad y un excesivo calentamiento en la parte central del equipo. No obstante, con este primitivo dispositivo era posible separar aminoácidos y péptidos en tres categorías: básicos, neutros y ácidos resultando imposible la posterior subdivisión dentro de cada uno de los grupos.

Años mas tarde, en 1954, Kolin publicó sus primeros trabajos sobre isoelectroenfoque aunque los gradientes de pH que empleaba eran muy cortos e inestables con el tiempo. Dos años después escribió un trabajo sobre el concepto de los números de transporte de los anfólitos, los compuestos responsables de generar el gradiente de pH. Vio que incluso en su punto isoelectrónico dichos compuestos tenían una conductancia apreciable si los valores de pK de los grupos disociados estaban próximos. Si se disponía de muchos de esos anfólitos, isoelectrónicos a varios pH, entonces en electrolisis deberían ser capaces de generar del ánodo al cátodo un suave gradiente de pH sin ningún mínimo de conductividad. Este era el concepto básico del isoelectroenfoque y los citados anfólitos se denominaron posteriormente «anfólitos transportadores». Pocos años mas tarde Svensson publicó la teoría básica del electroenfoque y en un trabajo posterior verificó experimentalmente que los anfólitos puros tenían conductividades que coincidían con sus predicciones teóricas. En el instituto Karolinska Svensson formó un grupo de trabajo en el que su objetivo prioritario fue encontrar los anfólitos adecuados.

Por hidrólisis parcial de la hemoglobina o de sangre total fue posible producir una serie de oligopéptidos que se utilizaron como anfólitos transportadores. A pesar de tratar con carbono y otros adsorbentes no fue posible obtener preparaciones de péptidos que no estuvieran coloreados. Sin embargo, el isoelectroenfoque demostraba que el color quedaba enfocado en una serie de bandas discretas. Animado por el resultado Vesterberg comenzó a probar de forma exhaustiva posibles métodos de síntesis de aminoácidos por unión de ácidos carboxílicos y aminas. En 1964 hirviendo a reflujo una mezcla de ácido acrílico y aminas polivalentes obtuvo una serie de anfólitos con valores de pK y puntos isoelectrónicos adecuados cuyas ventajas resultaron obvias desde el principio. Con el uso de esos nuevos anfólitos y utilizando un gradiente de sacarosa como medio estabili-

zante, era posible separar dos mioglobinas aunque las diferencias de sus puntos isoelectricos fuese tan solo de 0.05 unidades de pH siempre que se generaran gradientes de pH suficientemente estrechos. Se calculó que el poder de resolución teórico de la técnica era de 0.02 unidades de pH.

Tras 1970, Svensson, que cambió su nombre por el de Rilbe, diseñó un aparato para la rápida y adecuada preparación de electroenfoque en gradientes de densidad comercializado por LKB.

Otra contribución fundamental fue la aportada por Bjellquist al desarrollar el isoelectroenfoque en geles de poliacrilamida con cargas fijas inmovilizadas en la matriz del gel. Los gradientes de pH generados aseguraban la estabilidad del pH y permitían el enfoque de proteínas incluso en rangos de pH ultrafinos. Se llegó a reportar un poder de resolución de 0.001 unidades de pH.

Descrito, pues, el desarrollo de la electroforesis discontinua en geles de poliacrilamida en condiciones desnaturalizantes capaz de separar proteínas en función de su tamaño y las técnicas de electroenfoque capaces de hacer lo propio en función de la carga, ya tenemos la base para el posterior desarrollo de la electroforesis bidimensional de proteínas, punto de inicio de la actual proteómica.

Técnicas de proteómica

Las dos etapas fundamentales en proteómica son la separación de mezclas complejas de proteínas procedentes de un sistema biológico y la caracterización de los polipéptidos separados. En ambos casos se requiere de un potente apoyo bioinformático. Para el primer proceso se recurre normalmente a la electroforesis bidimensional, técnica capaz de separar miles de proteínas a la vez aunque recientemente se desarrolla la cromatografía bidimensional como posible alternativa. La caracterización de los polipéptidos se aborda mediante espectrometría de masas.

Separación de proteínas mediante Electroforesis Bidimensional

Hagamos también algo de historia. La primera aplicación de la electroforesis en geles de dos dimensiones se publicó en 1956 por Smithies y Poulik. Sin embargo los geles de almidón que utilizaron estaban lejos de ser los óptimos. Margolis y Kendrick recurrieron a geles de poliacrilamida preferentemente en gradiente, y obtuvieron significativas mejoras utilizando electroenfoque en un gel cilíndrico de poliacrilamida para

la separación mediante carga en la primera dimensión seguido de una electroforesis perpendicular en la segunda dimensión en un gel en placa como describieron Dale y Latner y Macko y Stegemann. En 1970, Stegemann introdujo el electroenfoque en poliacrilamida seguido de electroforesis en placa en presencia de SDS para aumentar la carga neta negativa de las proteínas y aprovechar la relación entre su tamaño y movilidad.

La separación de proteínas mediante electroforesis bidimensional de alta resolución se alcanzó definitivamente en 1975 gracias, fundamentalmente, al trabajo de Patric O'Farrell,

La electroforesis bidimensional en geles de poliacrilamida ha demostrado ser a fecha de hoy, la técnica mas poderosa para el análisis de proteínas procedentes de muestras biológicas complejas de muy diversa naturaleza. Inicialmente se requiere una primera etapa de preparación de la muestra en la que se solubilizan las proteínas. El método empleado debe ser compatible con el isoelectroenfoque así como no introducir modificaciones en las proteínas que pudieran afectar posteriormente a la espectrometría de masas. El dispositivo mas empleado para solubilizar proteínas recurre a lisar química o mecánicamente el tejido de interés. Frecuentemente es necesaria la centrifugación para separar las proteínas solubilizadas del material no solubilizado. Normalmente se utilizan detergentes zwitterionicos o detergentes no iónicos como el Nonidet P-40 (NP40) junto una alta concentración de urea.

El gran poder de resolución de la técnica surge al combinar dos separaciones basadas en principios químico-físicos diferentes: inicialmente las proteínas se separan mediante electroenfoque por carga al someterlas en condiciones «nativas» a la acción de un campo eléctrico en un gradiente de pH estabilizado físicamente en un gel de bajo porcentaje de acrilamida. De esta forma se logra separar las proteínas en función de sus puntos isoelectricos. El principio es simple y no requiere de excesiva explicación. Consideremos una columna en cuyo interior generamos un gradiente de pH que aumenta de forma continua desde el ánodo (parte inferior) al cátodo (parte superior). Con este dispositivo es evidente que al aplicar la corriente cada proteína se transporta hasta alcanzar el valor de pH que iguale a su punto isoelectrico, momento en que su carga neta se hará cero y se detendrá. El gran poder de separación de esta técnica nos lo da el dato de que es posible diferenciar proteínas que difieran tan solo en 0.01 unidades de pH en sus valores de pI. Posteriormente el fino cilindro de poliacrilamida que contiene las proteínas separadas por carga se desnaturaliza incubándolo durante unos pocos minutos en un exceso de SDS y se someten a un proceso de separación en gel de polia-

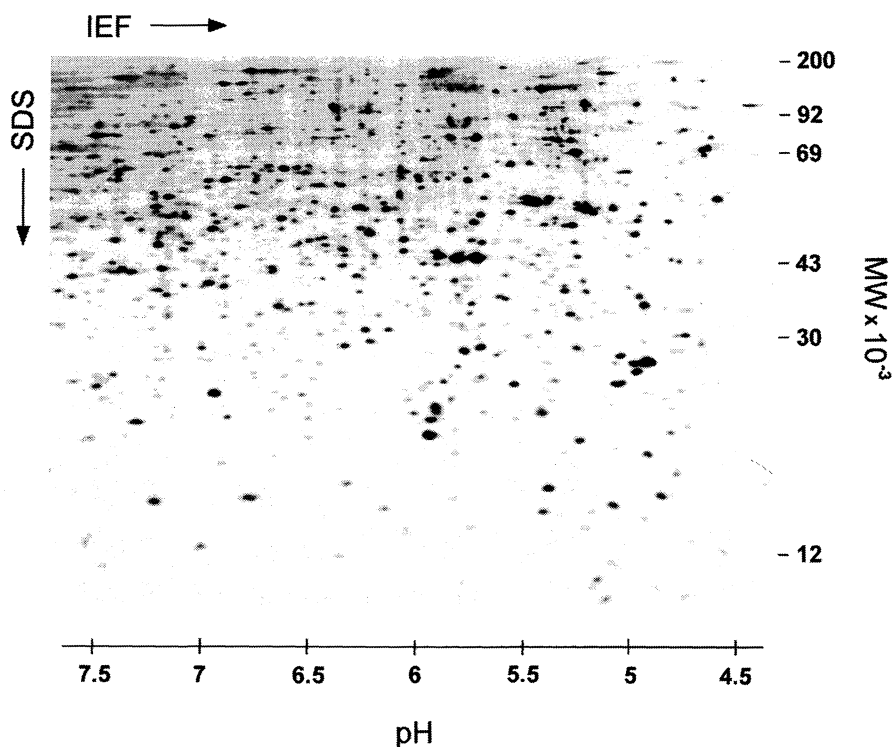


FIGURA 3. Electroforesis bidimensional de una mezcla de proteínas marcadas radiactivamente procedentes de un lisado de células de discos imaginales de ala de *Drosophila melanogaster*. En la primera dimensión las proteínas se han separado mediante electroenfoque en un gradiente de pH (IEF). Tras su desnaturalización se llevó a cabo la separación por tamaño en la segunda dimensión. El sistema permite resolver mas de 1200 polipéptidos.

crilamida en condiciones desnaturalizantes que las separara en función de sus diferencias de tamaño. De esta forma es posible resolver por encima de 2000 proteínas de forma rutinaria, dependiendo lógicamente del sistema biológico con el que se este trabajando y el método de detección utilizado (Figura 3).

Una pregunta obvia es ¿cómo se genera realmente un gradiente de pH para llevar a cabo la primera separación? Desde el año 1964 esto no ha sido problema para el investigador ya que a partir de esa fecha se comercializaron una serie de compuestos denominados Anfolitos, capaces de generar gradientes de pH dentro del rango de valores comprendido entre 9.5 y 4.5 que abrieron el camino de un amplio uso del electroenfoque, permitiendo la determinación exacta de los puntos isoelectricos de las proteínas en separaciones de alta resolución. Se trata de una serie de

compuestos de pequeño tamaño sintetizados a partir del ácido acrílico y una mezcla de polietilen-poliaminas en solución. Resultado de ello es una gran variedad de ácidos poliamino policarboxílicos con un amplio rango de puntos isoeléctricos que por su diferencia de carga son capaces de migrar en un campo eléctrico autogenerando el gradiente de pH al detenerse los propios compuestos en su pI y debido a su alta capacidad de tamponación. Además presentaban la propiedad de mostrar una apreciable capacidad de conducción de corriente incluso en su punto isoeléctrico. Uno de los problemas de estos compuestos es su inestabilidad a valores de pH básicos, lo que obligaba a llevar a cabo dos tipos de separación diferentes en función del tipo de proteínas (ácidos o básicas) se quisieran separar. Eso trajo como consecuencia una subdivisión de la electroforesis bidimensional que puede llevarse a cabo en condiciones de equilibrio o isoelectroenfoque (IEF) para separar proteínas ácidas con puntos isoeléctricos comprendidos entre 4.5 y 7.7 o bien en condiciones de no equilibrio (NEPHGE) para separar proteínas básicas cuyos puntos isoeléctricos se encuentren entre los valores de 7.5 y 9.0 aproximadamente.

Más recientemente, el protocolo se estableció en 1988, se ha comercializado otra variante que ha venido a simplificar y popularizar la técnica: es el empleo de gradientes de pH inmovilizados sobre acrilamida dispuestos para llevar a cabo la primera dimensión. Son las denominadas inmovilinas (IPGs) que han tratado de minimizar (con éxito relativo) los problemas de inestabilidad de los gradientes de pH en la región básica que hemos comentado antes, junto con una mayor estandarización del proceso. Indudablemente el empleo de las IPGs ha simplificado de forma muy considerable la técnica, aunque en opinión personal del autor cabe señalar que el nivel de resolución de los anfolitos es todavía superior al de las inmovilinas.

Visualización de las proteínas

Una vez separadas las proteínas mediante electroforesis bidimensional el siguiente problema a resolver es su visualización en el gel. Idealmente la intensidad de tinción debería correlacionarse con la cantidad de proteína y ser independiente de la naturaleza de la misma. Además, el método de tinción elegido no debe interferir con el posterior análisis mediante espectrometría de masas. Tradicionalmente las proteínas se han visualizado mediante marcaje radiactivo o por tinción cromofórica de los geles con Azul de Coomassie o, con mayor sensibilidad, con plata. Los métodos de detección radiactivos (por ejemplo autorradiografía de proteínas marcadas con [^{35}S] Metionina) son muy sensibles (comparados con la tin-

ción de plata) pero necesitan de largas exposiciones (hasta varias semanas) y tienen limitado rango dinámico si no se utilizan los métodos de fusión de imágenes que comentaremos mas adelante. La tinción con Coomassie es sencilla y muestra una buena reproducibilidad pero no es muy sensible y permite solamente la detección de los componentes mayoritarios de una muestra (normalmente no mas de 200-300 spots) presentes en cantidades en el rango de 100 ng. La tinción con plata puede llegar a ser 100 veces mas sensible que el Coomassie lo que la hace ideal para la detección de componentes minoritarios siendo todavía el método elegido para geles analíticos. El mayor problema de esta tinción es la baja reproducibilidad con ciertos protocolos, el bajo rango dinámico y el hecho de que ciertas proteínas se tiñen poco, negativamente o incluso no se tiñen. Actualmente hay una fuerte tendencia a utilizar sondas fluorescentes para la detección de proteínas en los geles. La tinción con fluoróforos es casi tan sensible como la tinción con plata pero tiene un mayor rango dinámico lineal (sobre tres órdenes de magnitud). Su mayor desventaja es el coste del equipo necesario para capturar la imagen.

Existen tres posibles aproximaciones a la detección de proteínas con sondas fluorescentes: (1) marcaje previo a la electroforesis, en el que las proteínas se acoplan a un fluoróforo antes de llevar a cabo la etapa de electroenfoque; (2) marcaje de las proteínas tras la primera dimensión y (3) marcaje tras la segunda dimensión. Actualmente la técnica mas popular es la tinción tras la electroforesis con SYPROTM Ruby. La tinción es no-covalente y se lleva a cabo en una única etapa que normalmente requiere pocas horas sin necesidad de destinción y que puede adaptarse fácilmente para utilizarla con instrumentos automáticos. El límite de detección es de 1-2 ng de proteína.

Aunque las proteínas abundantes son fácilmente detectables no debemos ignorar que en los estudios de proteomas normalmente solo somos capaces de detectar una pequeña fracción del total de proteínas expresadas por un sistema biológico. Las técnicas que tratan de visualizar simultáneamente un gran número de proteínas, como la electroforesis bidimensional que estamos considerando, pueden medir solamente un subset. A pesar del avance en la sensibilidad y los aspectos cuantitativos de la proteómica, un gran número de las proteínas están presentes a concentraciones por debajo de los límites de detección. Subsiste pues el problema de visualizar proteínas con un bajo nivel de expresión (un número de copias por célula comprendido entre 10-100) junto con el de las proteínas expresadas con alta frecuencia (denominadas proteínas «house-keeping», con mas de 10.000 copias por célula) que pueden llegar a enmascarar los componentes minoritarios. El prefraccionamiento, normalmente mediante

centrifugación, puede ser una solución a este problema, ya que combinado con la alta capacidad de carga de los geles preparativos parece que pueden llegar a poner de manifiesto incluso las proteínas minoritarias.

Señalemos por último que las moléculas de proteína con pesos moleculares muy altos (>200.000) o muy bajos (<8.000) tampoco son fáciles de detectar y las restricciones de solubilidad hacen que sea difícil el estudio de las proteínas de membrana.

Captura y análisis de imágenes

Los patrones de proteínas generados a partir de la electroforesis bidimensional normalmente son complejos y es recomendable analizarlos mediante computación. De esta forma es posible crear bancos de datos de proteínas, algunos de los cuales están disponibles en la red (<http://us.expasy.org>).

Las etapas fundamentales en el análisis de imágenes mediante computador son la adquisición de datos, la reducción ruido, la detección y cuantificación de spots, el «emparejamiento» de los patrones y la construcción del banco de datos. Las imágenes se capturan «escaneando» los geles teñidos con Coomassie o con plata o las películas sensibles a rayos X.

Antes de la detección de «spots» se elimina el ruido en una etapa totalmente automatizada tras la cual deben detectarse y cuantificarse las manchas de proteína. Aunque esto parece simple, la detección correcta de los spots individuales no es fácil de llevar a cabo porque no todas las proteínas se resuelven como spots discretos. Por tanto, si bien es cierto que la detección y la cuantificación de spots se llevan a cabo automáticamente no existe un método para corregir de forma inequívoca los spots ni para analizar aquellos que solapen. Consecuentemente, los spots tienen que editarse manualmente en un proceso que lleva mucho tiempo, y que depende del número de spots y de la calidad de la separación electroforética.

Una parte fundamental de los estudios de proteómica, especialmente en el contexto de las investigaciones patológicas, es la comparación de los patrones de expresión de proteínas. Para llevar a cabo dicho análisis es necesario utilizar softwares especiales que sean capaces de comparar todos los spots de un gel con sus contrapartes (caso de que existan) en los otros geles. La lógica de operación es la siguiente: Uno de los geles (normalmente el de mayor calidad) se designa como referencia. A continuación, al menos en la mayoría de los programas de análisis de imágenes,

hay que seleccionar una serie de spots de referencia («landmark») que deben estar distribuidos de manera uniforme sobre todo el gel y presentes en todas las muestras. Partiendo de esas referencias el programa ya es capaz de comparar el resto de spots, comenzando por los vecinos mas próximos a las referencias y avanzando hacia los mas distantes. Tras el emparejamiento hay que comprobar nuevamente si las operaciones se han hecho correctamente y corregir las que hayan sido defectuosas. De esta forma es posible comparar los múltiples patrones proteicos de los geles bidimensionales de una serie de experimentos.

Los softwares deben ser capaces de comparar múltiples grupos de geles así como admitir información adicional que quiera introducir el usuario. La información guardada en el banco de datos puede utilizarse para revelar diferencias cualitativas o cuantitativas entre muestras individuales detectando proteínas expresadas diferencialmente en situaciones que se alejen del control (patológicas, mutantes, etc..).

El primer sistema de análisis automático que se comercializó fue el PDQUEST, desarrollado por James Garrels en Cold Spring Harbor en 1985. Fue el mas elaborado y perfecto ya que además de las operaciones ya señaladas el PDQUEST permitía analizar con extraordinario detalle patrones de expresión proteicos procedentes de películas autorradiográficas al posibilitar fusionar las imágenes obtenidas de diferentes exposiciones de un mismo gel generando una imagen virtual con un rango dinámico superior a cuatro órdenes de magnitud. Curiosamente las actuales versiones comerciales de este sistema han eliminado esta operación con lo que ha disminuido considerablemente su utilidad. Otros sistemas de análisis similares son el ELISE, el HERMES o el MELANIE que utilizan estrategias análogas para la comparación de geles.

Identificación de Proteínas

Visualización no es lo mismo que identificación y hasta hace relativamente poco este era el cuello de botella de la proteómica. El peso molecular y el punto isoelectrico de una proteína, información que proporciona la electroforesis bidimensional, es insuficiente para su identificación y realmente el punto clave para aprovechar la información obtenida en los geles reside en la capacidad para identificar (y caracterizar) las proteínas de interés. En el pasado la identificación estaba limitada al wester blot en los casos en los que se disponía de un buen anticuerpo o, si las proteínas eran muy abundantes, se hacía uso de los secuenciadores automáticos. En este último caso se podía determinar directamente la secuencia

del amino terminal de la proteína intacta (caso de que no estuviese bloqueado) o bien se podía recurrir a la digestión enzimática (en el propio gel o tras transferir a una membrana) seguida de la separación de péptidos por HPLC para determinar posteriormente la secuencia interna de alguno de ellos mediante degradación de Edman. La limitación de ambos procesos era la cantidad de proteína/péptido necesaria para proporcionar una secuencia sin problemas. Con estos métodos se requiere al menos de unos pocos picomoles de proteína para tener éxito con el análisis.

Dos hechos han dado un vuelco a esta situación en los últimos años: el desarrollo de técnicas de espectrometría de masas de alta resolución junto con el crecimiento de los bancos de datos de secuencias.

Los espectrometros de masas consisten en una fuente para generar iones a partir de una muestra y un analizador capaz de separar y detectar dichos iones de acuerdo con sus masas. Para proteómica existen dos fuentes de ionización que se utilizan habitualmente: la ionización-desorción asistida por laser sobre una matriz (MALDI) y la ionización por electrospray (ESI). Por su parte los analizadores pueden ir de los mas sencillos de tiempo de vuelo (TOF) a los mas complejos (por ejemplo Fourier-transform ion-cyclotron resonance, o FTICR). La salida, sin embargo, es siempre la misma: la determinación exacta de la masa de los iones moleculares y de sus fragmentos.

A esto hay que añadir el rápido crecimiento de los bancos de datos: miles de millones de bases de secuencias genómicas de multitud de organismos están ahora disponibles en la red. Para el punto que aquí nos ocupa digamos que se han desarrollado softwares y buscadores capaces de conectar los bancos de datos de genomas y proteínas con la información emanada de los espectros de masas.

Uno de los métodos mas utilizados para identificar proteínas es recurrir al análisis de su «huella peptídica» (Figura 4). Dado que la masa de una proteína es un dato insuficiente para identificarla de manera inequívoca lo que se hace es convertir las proteínas en péptidos mas cortos. Para ello se recorta el trocito de gel bidimensional que contiene la proteína que queremos identificar y se digiere con una enzima proteolítica, como tripsina, que rompe específicamente a derecha de aminoácidos lisina y arginina, generando una serie de fragmentos trípticos únicos y específicos para cada proteína. A continuación se recurre a la espectrometría para determinar las masas de esos péptidos derivados de la digestión proteolítica (Figura 4a). La técnica mas habitual para hacer esto es el MALDI-TOF (matriz-assisted laser desorption ionization time-of-fly). El sistema es muy sensible y puede dar información sobre el peso molecular a nivel de attomoles (10^{-14} g para una molécula de 10 kD).

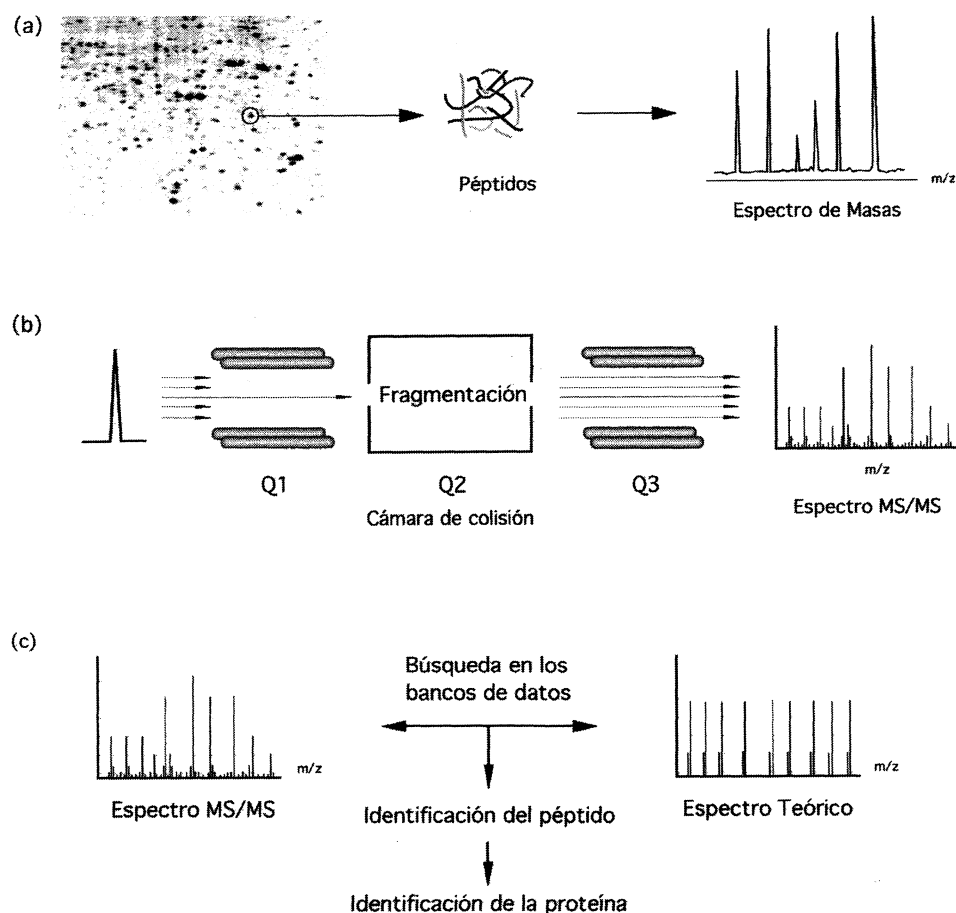


FIGURA 4 Análisis proteómico mediante electroforesis bidimensional y espectrometría de masas. El método consiste en tres etapas que están íntimamente integradas (a) Las proteínas se separan por electroforesis bidimensional y los spots teñidos se extraen y se someten a una digestión con tripsina en el propio gel. Los péptidos resultantes se analizan por espectrometría de masas para determinar su «huella peptídica» que puede conducir a la identificación de la proteína. (b) Es posible seleccionar uno de los péptidos separados en el espectro de masas haciéndolo pasar a través de un cuadrupolo (Q1) para fragmentarlo en una cámara de colisión (Q2) y recoger y analizar los nuevos iones en un tercer cuadrupolo (Q3). (c) El espectro generado a partir del ión seleccionado normalmente contiene suficiente información para deducir su secuencia e identificar el péptido y por tanto la proteína de la que se originó.

Esas masas obtenidas se utilizan entonces para hacer una búsqueda en los bancos de datos de masas de péptidos, que se han generado *in silico*

(utilizando roturas específicas) a partir de secuencias de aminoácidos, incluyendo las deducidas a partir de secuencias de DNA. Esta estrategia, aunque utilizando diferentes algoritmos, fue desarrollada por diversos grupos independientes en 1993. Los programas de análisis tratan de ajustar los datos de las masas del usuario a una de las secuencias de proteína existentes en el banco de datos y sugieren la identidad de la proteína. Teniendo en cuenta la precisión en la medida de las masas actualmente alcanzables (mejores de 10 ppm) esta técnica sola puede ser suficiente para identificar proteínas a partir de genomas totalmente secuenciados ya que es posible calcular de forma precisa las masas teóricas de todos los péptidos tripticos de todas las fases de lectura abiertas posibles. Para una identificación correcta deben coincidir al menos cuatro o cinco péptidos experimentales con los teóricos y se necesita cubrir un mínimo del 25 % de la secuencia. Debemos insistir que esta forma de identificación de proteínas solo es posible si la proteína que se está analizando se encuentra registrada en el banco de datos.

Si no se dispone de secuencias en los bancos de datos los espectrómetros de masas pueden generar información no solo sobre las masas sino sobre las propias secuencias de alguno de los péptidos digeridos. Para ello, en lugar de utilizar un espectrómetro de masas simple como el MALDI-TOF que mide solamente la masa, se puede utilizar un espectrómetro de masas en tandem con un cuadrupolo híbrido TOF capaz de generar la información de secuencia necesaria. La forma de proceder es la siguiente. Inicialmente se obtiene el espectro general como hemos descrito anteriormente y sobre el se selecciona el ión cuya secuencia queramos determinar. Seguidamente se hace un nuevo proceso en el que mediante un filtro de masas tan solo el ión seleccionado es capaz de atravesar el primer cuadrupolo. A continuación se le fragmenta por colisión con un gas analizando los fragmentos generados en un segundo analizador TOF que mide sus masas (Figura 4b). En una última etapa el espectro de fragmentación obtenido se interpreta para deducir la secuencia de aminoácidos del péptido y, a partir de ella, eventualmente identificar la proteína (Figura 4c). Caso de que el péptido encontrado no estuviese en ningún banco de datos puede utilizarse para, a partir de la secuencia deducida, diseñar un oligo degenerado que se utilice posteriormente como sonda para acceder al gen que codifica la proteína problema.

El MALDI MS es indudablemente uno de los pilares fundamentales de la proteómica. La instrumentación es robusta, relativamente barata y admite un alto grado de automatización. Con un tiempo de análisis por muestra de unos pocos segundos existe un gran potencial para analizar varios miles de muestras por semana si se desarrollan el automatismo

necesario para cortar los spots y hacer la proteólisis de los péptidos. Los métodos de MALDI-TOF son capaces de detectar bajos niveles de proteína. Actualmente, desde un punto de vista realista, se necesitan varios cientos de femtomoles de proteína por spot en un gel para poder llevar a cabo una caracterización siendo necesario cantidades del orden picomolar para el análisis de modificaciones post-traduccionales. Esto nos lleva a que si consideramos unas 1000 copias de proteína por célula necesitaríamos del orden de 10^8 - 10^9 células para hacer una identificación.

Sin embargo el MALDI MS tiene algunas limitaciones. En primer lugar la ionización es selectiva y no es cuantitativa. En una mezcla equimolar de péptidos procedentes de un digerido de una proteína algunos de ellos no se detectarán y habrá una amplia variación en intensidad de señal para los restantes. En segundo lugar, a medida que decrece la cantidad de proteína en el gel lógicamente también lo hace el número de péptidos y a muy bajos niveles de proteína solo es posible observar unos pocos péptidos. La masa de un único péptido por sí es insuficiente para definir la secuencia. Así, es imposible la identificación de una proteína con garantías a partir de unos pocos péptidos y esta situación se da con frecuencia cuando se dispone de poca cantidad de proteína.

El electrospray (ESI) forma el segundo pilar para la espectrometría de masas en proteómica aunque el mecanismo de formación de iones en esta modalidad es bastante diferente del MALDI. Basándose en estudios teóricos se demostró que era posible obtener una ionización eficiente mediante un spray, haciendo pasar el digerido de péptidos a través de un fino tubo capilar a bajas velocidades de flujo (típicamente, 25 nl min^{-1}). Una muestra de 1 ml podría estar pulverizándose durante una hora en una fuente de nano electrospray (nESI) permitiendo muchos experimentos de secuenciación detallados de un ión precursor tras otro, o hacer barridos de precursor y producto para estudios de modificaciones post-traduccionales. La técnica permite el acoplamiento de sistemas de separación «on-line» tales como la cromatografía líquida de alta presión (HPLC) y la electroforesis capilar al espectrómetro de masas, permitiendo el análisis de mezclas muy complejas. En ESI, los iones precursores procedentes de una digestión trípica normalmente tienen dos cargas, una de las cuales suele estar deslocalizada y ayuda a inducir fragmentación. La selección de iones y la fragmentación también se puede hacer en un analizador simple con trampa iónica y FTICR. Los principios de la MS en tandem son los mismos.

El sistema es muy robusto y se ha adaptado a la mayoría de los espectrómetros comerciales. Quizá la última fuente de nESI es la inyectable fine-ionization source (JaFIS) que es el resultado de una muy

fin a optimización. A velocidades de flujo muy bajas ($1\text{-}2\text{ nl min}^{-1}$) es posible mantener un flujo estable durante 20 horas.

Proteómica cuantitativa

Resuelto el problema de la identificación de proteínas uno de los objetivos siguientes se ha enfocado en el desarrollo de la proteómica cuantitativa, que permita obtener información sobre la dinámica de un sistema biológico. Son muchos los casos en los que resulta fundamental conocer de manera precisa y cuantitativa, los cambios que experimenta una muestra ante una determinada perturbación o cambio fisiológico espontáneo o inducido. Sería, en cierta medida, un sistema complementario a los arrays de cDNA que tratan de cuantificar los cambios de expresión génica a nivel de RNA.

La aproximación experimental ha recurrido al uso de isótopos estables. La idea es que para llevar a cabo análisis de un analito(s) concreto se mezcle con una muestra idéntica, pero marcada con isótopos pesados estables (p. ej: ^2H , ^{13}C , ^{15}N). Con ello se resuelve el problema del grado de ionización de los péptidos que queramos analizar ya que va a ser el mismo independientemente de que procedan de la muestra ligera o de la pesada. De esta manera es posible llevar a cabo cuantificaciones si analizamos una mezcla de dos poblaciones, una de las cuales sirve como referencia al contener las mismas proteínas que la otra pero con abundancia diferente y marcadas con isótopos pesados estables. Teóricamente, tras la combinación de muestras, todos los péptidos existirán como parejas de analitos de idéntica secuencia pero con diferentes masa. Cada pareja de péptidos tendrán las mismas propiedades fisicoquímicas y se comportarán de forma idéntica ante cualquier etapa de aislamiento o separación por lo que la relación entre las intensidades de las masas mayor y menor de los componentes de esas parejas de picos en el espectro proporcionará una medida precisa de la abundancia de los péptidos (y por extensión de las proteínas) en la muestra original.

Esta es la teoría. El problema es cómo se incorporan los isótopos estables y como se analizan posteriormente. En este sentido existen dos alternativas para llevar a cabo el marcaje que son, en resumen, las siguientes:

1. *Marcaje previo a la extracción:* El sistema debe ser capaz de incorporar metabólicamente el precursor. Imaginemos un cultivo de un microorganismo. Podemos hacerlo crecer en un medio que contenga una distribución natural de nitrógeno (^{14}N 99,6 %; ^{15}N 0,4 %) o bien en un medio

enriquecido en ^{15}N (>96 %). Tras crecer los cultivos en los medios anteriormente señalados y en las condiciones experimentales que elijamos para cada uno se reúnen ambos en un «pool» celular común que se procesa como estimemos mas conveniente. Por ejemplo podemos extraer las proteínas de interés, separarlas por HPLC y posteriormente por SDS-PAGE. Podemos digerir los spots seleccionados en el gel y analizar los fragmentos mediante espectrometría de masas. La huella peptídica permitirá la identificación de la proteína como ya hemos descrito anteriormente, pero además nos encontraremos que cada átomo de ^{15}N incorporado desplazará la masa de cualquier péptido hacia una unidad de masa superior originando «parejas de picos» para cada péptido. Esta estrategia permitirá la cuantificación de las dos muestras que queríamos comparar con un alto grado de precisión.

El método tiene sus limitaciones. En primer lugar no permite el análisis de las proteínas directamente a partir del tejido. En segundo lugar, el medio enriquecido en isótopos estables puede afectar de por si a la proliferación, y por tanto introducir artefactos a nivel de las proteínas. En tercer lugar los medios enriquecidos en isótopos estables son caros y para cultivo de células de organismos superiores son imposibles de obtener. En cuarto lugar, el aumento en masas nominal producido por la incorporación del isótopo estable no se conoce hasta que se determina la secuencia, lo cual puede confundir la búsqueda en los bancos de datos.

2. *Marcaje posterior a la extracción*: Estos métodos responden a la siglas de ICAT (Isotope-Coded-Affinity-Tags) y su principio se esquematiza en la Figura 5. En esta ocasión la incorporación de los isótopos estables tiene lugar después del aislamiento de la proteína de interés mediante la alquilación selectiva de los restos de cisteína con reactivos «ligeros» o «pesados» (Figura 5a) tras lo cual se mezclan las dos proteínas que se quieran cuantificar. A partir de aquí se puede introducir cualquier etapa encaminada a enriquecer las proteínas minoritarias de interés o para reducir la complejidad de la muestra caso de que sea necesario. Las cantidades relativas de las dos muestras se mantendrán invariables. Antes del análisis final la mezcla de proteínas se digiere con tripsina y a continuación se pasa por una columna de agarosa-avidina. Como el marcaje con ICAT posee una etiqueta de biotina (Figura 5a) es posible aislar de manera selectiva los péptidos marcados con ICAT que contienen cisteína para su posterior análisis por masas. La relación de intensidades de los iones de las parejas de péptidos-ICAT que aparezcan permiten su cuantificación. El posterior análisis mediante MS/MS permite identificar la proteína (Figura 5b).

Esta estrategia de ICAT tiene evidentes ventajas. En primer lugar el método no necesita de incorporación metabólica, por lo que es compatible

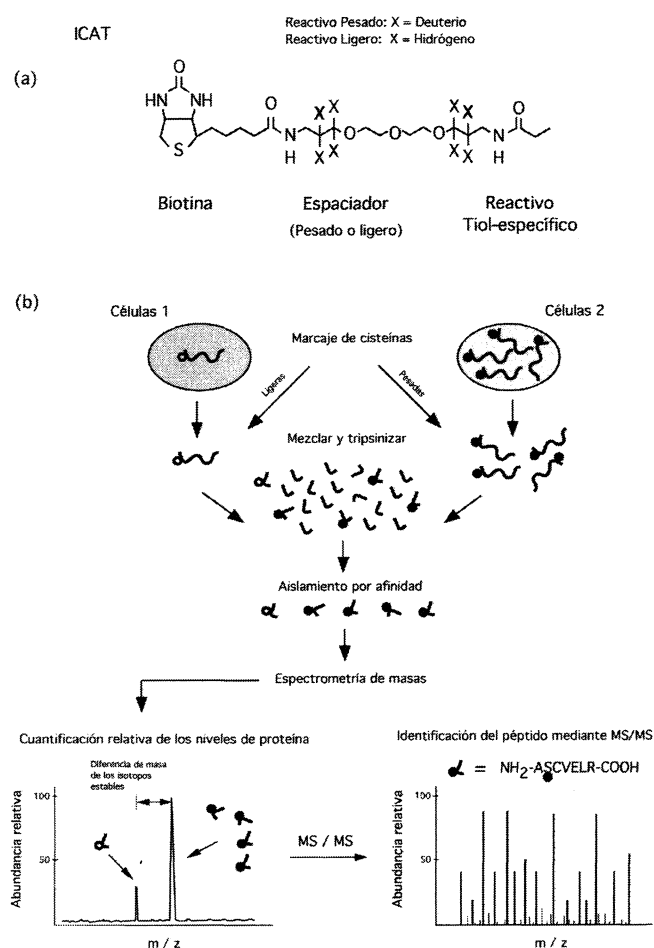


FIGURA 5. Estrategia para la cuantificación de expresión diferencial de proteínas mediante ICAT (Isotope-Coded Affinity Tag). (a) Estructura del reactivo ICAT. Consta de tres elementos: una etiqueta de afinidad (biotina); un espaciador que puede incorporar isótopos estables; y un grupo reactivo que es específico para los grupos tiólicos (por ejemplo para cisteínas). El reactivo existe en dos formas: pesado (conteniendo ocho átomos de deuterio) y ligero (conteniendo ocho átomos de hidrógeno). (b) Estrategia del ICAT. Se muestra el análisis para una única proteína pero es aplicable a un lisado celular total. Se extrae la proteína de los dos estados celulares diferentes, se desnaturaliza, reduce y se marca en las cisteínas con el reactivo ICAT ligero o pesado. Las muestras se mezclan y se digieren con tripsina. Seguidamente se aíslan los péptidos marcados con ICAT mediante cromatografía de afinidad con biotina y a continuación se analizan por HPLC acoplada a espectrometría de masas en tandem. La relación de intensidades para una pareja de iones marcados con ICAT cuantifica la abundancia relativa de sus proteínas parentales en el estado celular original. Además el espectro en tandem revela la secuencia del péptido e identifica de manera inequívoca la proteína.

con cualquier cantidad de proteína extraída de cualquier fluido fisiológico, células o tejidos bajo cualquier condición de crecimiento. La reacción de alquilación es altamente específica y tiene lugar en presencia de sales, detergentes o agentes estabilizantes (por ejemplo SDS, urea, HCl-guanidina). En tercer lugar la complejidad de la mezcla de péptidos a ser analizada se reduce por el aislamiento exclusivo de los péptidos que contienen cisteína. En cuarto lugar la estrategia de ICAT permite casi cualquier tipo de fraccionamiento bioquímico, inmunológico o físico, que lo hace compatible con el análisis de proteínas poco abundantes.

Existen dos inconvenientes en el método. En primer lugar el marcaje con el ICAT introduce una modificación relativamente grande (500 Da) que permanece en cada péptido a lo largo de todo el análisis. Esto puede complicar las búsquedas en los bancos de datos especialmente para péptidos pequeños (< 7 aminoácidos). En segundo lugar el método falla para proteínas que no contengan cisteínas. Sin embargo, solo una pequeña fracción de las proteínas carece de cisteínas y además es posible sintetizar otros reactivos ICAT con grupos específicos distintos del tiol.

Métodos alternativos de análisis en proteómica: cromatografía multidimensional

Hemos visto que la separación de proteínas mediante electroforesis bidimensional presenta algunas limitaciones. El análisis de proteomas a gran escala requiere extraer, digerir y analizar individualmente cada uno de los spots de un gel bidimensional en un proceso que consume mucho tiempo. Además, aparte de la limitada capacidad de carga de los geles bidimensionales y los límites de los métodos de tinción, la técnica tiene normalmente un insuficiente rango dinámico para el análisis completo de proteomas. El rango dinámico de un sistema proteómico puede definirse como el número de copias por célula de la proteína mas abundante identificada dividida por el número de copias por célula de la proteína menos abundante identificada en el sistema. Todas esas limitaciones de la electroforesis bidimensional como método de separación ha animado a la búsqueda y desarrollo de métodos alternativos para el análisis de péptidos y proteínas a gran escala.

Algunas alternativas de separación recurren a métodos cromatográficos mono y bidimensionales que utilizan cromatografía líquida de alta resolución (HPLC), isoelectroenfoque capilar (CIEF), electroforesis capilar (CE) o cromatografía microcapilar. Como en el caso de la electroforesis

bidimensional, la espectrometría de masas sigue siendo el método elegido para la identificación de las proteínas resueltas por los métodos de separación líquidos.

Lo habitual es separar las proteínas en una primera dimensión por tamaño (mediante una cromatografía de exclusión) o por carga (mediante intercambio iónico) y en la segunda dimensión hacer uso de HPLC en fase reversa. Este sistema produce cromatogramas bidimensionales que son análogos a los geles bidimensionales con spots discretos separados en los ejes X-Y. Aunque actualmente estos sistemas están lejos del poder de resolución de la electroforesis bidimensional existe un considerable margen de mejora. El potencial del método parece superior y mas rápido que el de los geles y puede alcanzar un mayor grado de automatización.

Esta aproximación parece prometedora para pequeños genomas bacterianos y puede ser eficaz también para complejos de proteína. Como ejemplo de cromatografía bidimensional seguida de espectrometría de masas en tandem señalemos que las proteínas del periplasma de *E. coli* se fraccionaron mediante un intercambiador aniónico fuerte en HPLC y alícuotas de cada una de las 20 fracciones se digirieron con tripsina y se analizaron por HPLC en una microcolumna de fase reversa. Los péptidos eluidos de la microcolumna se aplicaron directamente en un espectrómetro de masas en tandem que permitió la identificación de 80 proteínas de forma totalmente automática.

Bibliografía

- Methods in Molecular Biology: Volume 112. 2-D Proteome análisis protocols. Edited by ANDREW J. LINK. Humana Press. Totowa, New Jersey. 1999.
- Proteome Research: Two-dimensional gel electrophoresis and identification methods. Edited by T. RABILLOU. Springer. 2000.
- Proteomics. From protein séquense to function. Edited by S. R. PENNINGTON and M. J. DUNN. BIOS Scientific Publishers Ltd. 2001
- Proteome Research: Mass spectrometry. Edited by PETER JAMES. Springer. 2001.
- Protein sequencing and identification using tandem mass spectrometry. MICHAEL KINTER and NICHOLAS E. SHERMAN. Wiley-Interscience. 2000.